Multiscale spatiotemporal heterogeneity analysis of bike-sharing system's self-loop phenomenon: Evidence from Shanghai

Yichen Wang

Division of Logistics and Transportation, Tsinghua University, 100084 Beijing, China

Email: wang-yc22@mails.tsinghua.edu.cn

Qing Yu

School of Urban planning and design, Beijing University, 100084 Beijing, China

Email: yuq@pku.edu.cn

Yancun Song

Polytechnic Institute & Institute of Intelligent Transportation Systems, Zhejiang University, Hangzhou 310058, China

Email: 22260191@zju.edu.cn

Abstract: Bike-sharing is an environmentally friendly shared mobility mode, but its self-loop phenomenon, where bikes are returned to the same station after several time usage, significantly impacts equity in accessing its services. Therefore, this study conducts a multiscale analysis with a spatial autoregressive model and double machine learning framework to assess socioeconomic features and geospatial location's impact on the self-loop phenomenon at metro stations and street scales. The results reveal that bike-sharing self-loop intensity exhibits significant spatial lag effect at street scale and is positively associated with residential land use. Marginal treatment effects of residential land use is higher on streets with middle-aged residents, high fixed employment, and low car ownership. The multimodal public transit condition reveals significant positive marginal treatment effects at both scales. To enhance bike-sharing cooperation, we advocate augmenting bicycle availability in areas with high metro usage and low bus coverage, alongside implementing adaptable redistribution strategies.

Keywords: Bike-sharing, self-loop, land use, emerging mobility, shared mobility, double machine learning

1. Introduction

The bike-sharing systems serve as an environmentally friendly emerging mobility service option, playing increasingly important roles as a vital substitute for multimodal urban public transit systems (Hosford et al., 2024, Diao et al., 2023). Since their introduction in China in 2016, bike-sharing has led to notable environmental and economic benefits (Sun and Ertz, 2022). As shown in Fig. 1, the self-loop phenomenon in the bike-sharing system refers to the observed spatial returning behavior of bikes to their original starting regions after several trips in the bike mobility chain (Song et al., 2024). This unique phenomenon in daily operations can be utilized to address negative sentiments from the public's concern about the improper parking management practice and supply-demand spatiotemporal imbalance in bike-sharing systems' daily operation, if the mechanisms behind its formation can be well understood (Zhao et al., 2022, Shi et al., 2024, Li et al., 2024b).

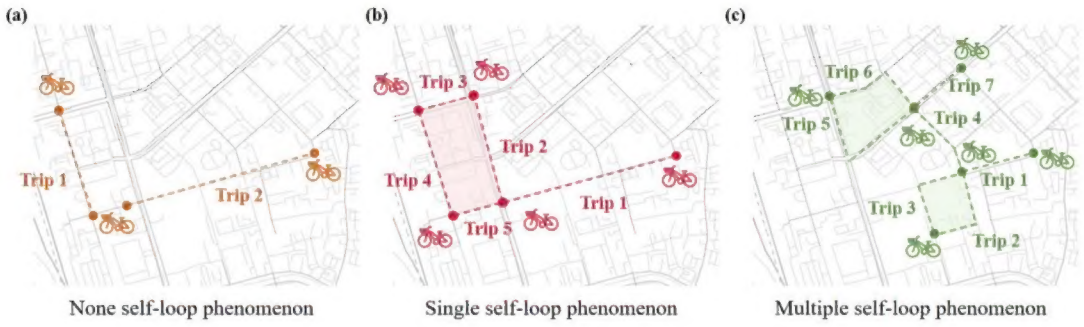

Fig. 1. Illustration of the self-loop phenomenon in the bike-sharing system. **(a)** None self-loop phenomenon. **(b)** Single self-loop phenomenon. **(c)** Multiple self-loop phenomenon.

Understanding the self-loop phenomenon is crucial for optimizing bike-sharing operations and enhancing their integration with other urban public transit modes (Yu et al., 2024). Many previous studies have highlighted the significant influence of land use patterns on bike-sharing usage (Kim and Cho, 2021, Jaber and Csonka, 2023). Land use and built environment characteristics such as residential density, commercial areas, and proximity to public transit stations affect bike-sharing demand's spatial and

temporal distribution (Yu et al., 2020, Pekdemir et al., 2024). As a substitute for urban rail transit and bus transit, bike-sharing systems often serve as first-mile and last-mile solutions complementing mass transit, which means transit passenger flow features are vital in shaping bike-sharing patterns (Yan and Chen, 2024, Chu et al., 2021). While previous studies have examined overall usage patterns and the integration of bike-sharing with other transit modes, there is still limited research in understanding the mechanisms behind the formation of self-loops in bike-sharing systems (Song et al., 2024, Xin et al., 2023), particularly about land use diversity and the positioning of bike-sharing within the multi-modal transportation network, remain unexplored.

Therefore, this study contributes to the existing research gaps by exploring factors influencing the bike-sharing self-loop phenomenon. First, we focus on quantifying the impact of socioeconomic attributes, land use patterns, and multimodal public transit conditions on bike-sharing self-loops. By adopting a multiscale framework that examines both metro station and street scales, we capture the heterogeneity of influencing mechanisms in different urban contexts. Second, a double machine learning framework is employed to capture the underlying treatment effects across different urban contexts. And from a practical perspective, we provide strategic recommendations for enhancing bike-sharing operation efficiency and integrating with multimodal public transit networks, tailored to different geospatial location and socioeconomic scenarios.

The remainder of this study is structured as follows. Section 2 reviews the impacts of land use and the relationship between bike-sharing and multimodal public transit systems. Section 3 details the methodology, outlining the spatial econometric and causal inference approaches utilized for the analysis. Section 4 introduces the study area, data source, bike-sharing self-loop intensity calculation process, and

empirical analysis, focusing on spatial heterogeneity and causal relationships at the metro station and street scales. Section 5 discusses the implications of the findings for urban planning and bike-sharing operations and provides policy recommendations. Section 6 concludes with limitations and future work.

2. Literature review

2.1 Land use impact on bike-sharing spatial demand patterns

Bike-sharing systems have evolved into an essential component of urban mobility, particularly offering a sustainable solution for short-distance trips (Li et al., 2024a). Initially introduced to reduce congestion and environmental pollution, bike-sharing has been a practical approach to improving the efficiency of city transit networks (Huang and Xu, 2023). Integrated seamlessly with public transportation, these systems enhance the accessibility of urban transit by providing reliable linkages to major metro stations and bus stops (Guo and He, 2020, Zhou et al., 2024). This integration not only optimizes public transit usage but also encourages a modal shift towards more active and low-emission transport options (Saltykova et al., 2022).

The spatial patterns of bike-sharing usage are being well-understood from multiple perspectives, with a significant association between bike usage and built environment, socioeconomic attributes, and transportation infrastructure (Mix et al., 2022, Wang et al., 2024). Central business districts, transportation hubs, and residential zones are frequently identified as hotspots for bike-sharing activity (Xing et al., 2020). These areas tend to experience high bike usage levels due to their accessibility, functional diversity, and high population densities (Eren and Uz, 2020). The spatial concentration of demand in these hotspots reflects the importance of land use in shaping bike-sharing utilization (Li et al., 2021). For instance, areas with a mix of residential and commercial functions foster more consistent usage throughout the

day, as different trip purposes (e.g., commuting, shopping, and entertainment) create overlapping and frequent demand (Xing et al., 2020). Understanding these spatial patterns provides insights into where bike-sharing fleets should be positioned to effectively serve local travel demand (Gao et al., 2021).

One persistent operational challenge for bike-sharing systems is maintaining the balance between supply and demand across different stations (Cheng et al., 2021). Imbalances such as bike shortages are common in busy urban locations (Zheng et al., 2024). To alleviate these issues, rebalancing strategies are frequently employed, often utilizing prediction models to distribute bikes according to anticipated demand patterns (Li et al., 2023). However, in some instances, a natural equilibrium is achieved without intervention, a phenomenon defined as "self-looping," where bikes are often returned to their origin stations (Song et al., 2024). This inherent self-regulation not only assists in maintaining local availability but also reduces the reliance on costly rebalancing operations.

Land use features are fundamental in influencing bike-sharing usage patterns and understanding these influences is vital for enhancing system performance (Ji et al., 2023). Residential areas generally exhibit peaks in bike use during morning and evening commutes, whereas commercial districts show steady demand throughout the workday (Ji et al., 2023). Such diversity in land use provides an environment where frequent, short-distance cycling is convenient, ultimately leading to higher overall bike-sharing utilization (Guo and He, 2020). As a consequence, exploring the relationship between land use patterns and bike-sharing demand offers a valuable understanding of how bike-sharing operations could be improved in diverse urban contexts (Mix et al., 2022). The literature review of land use impact on bike-sharing spatial demand patterns is shown in Table 1.

Table 1. Literature review of land use impact on bike-sharing spatial demand patterns.

Author/Year	Land use characteristics	Method	Study case	Major takeaways
Gao et al. (2021)	Population density, commercial/industrial land use	MLR, AGWR	Shanghai, China	Built environment influences bike-sharing distance decay, varying across urban contexts
Guo and He (2020)	Mixed-use, residential, industrial	Negative binomial regression	Shenzhen, China	Mixed land use positively influences bike-sharing integration with metro
Guo and He (2021)	Perceived built environments	Path analysis, perception survey	Shenzhen, China	The perceived built environment impacts dockless bike-sharing and metro integration more significantly than objective measures
Jaber and Csonka (2023)	Financial services, educational, and residential areas	Multinomial logit models	Hohhot, China	Commuting, dining, and service activities influence bike-sharing behavior
Ji et al. (2023)	Land-use diversity, Residential POI, Commercial POI	Generalized additive mixed modeling	Xi'an, China	Land use density shows an inverted U-shaped relationship with bike-share arrivals and departures.
Li et al. (2024a)	Mixed residential, commercial, and industrial	Spatial probit regression	New York City, USA	Land use and socio-demographic factors significantly impact bike-sharing usage resilience
Mix et al. (2022)	Residential areas, office land use	Location optimization models	Santiago de Chile	Built environment and accessibility variables could be used to optimize bike-sharing system station locations
Pekdemir et al. (2024)	Residential areas, educational areas, public parks, commercial areas,	OLS, Partial least squares regression	Izmir, Türkiye	Transport hubs have a significant impact on bike-sharing trips
Wang et al. (2021)	Commercial POI, Service POI	Regression with spatially varying coefficients	Montreal, Canada	Commercial POIs have varying effects depending on spatial context.
Zhao et al. (2020)	Residence, work POI	Ensemble clustering	San Francisco Bay Area, USA	Mixed land-use areas have significantly higher bike-sharing activity

Note: Adaptative Geographically Weighted Regression (AGWR), Multiple Linear Regression (MLR), Ordinary Least Squares (OLS)

2.2 Relationship between bike-sharing and multimodal public transit systems

The relationship between bike-sharing systems within multimodal public transit networks is primarily reflected in their capacity to enhance the efficiency of public transportation transfers (Caggiani et al., 2020). The inherent flexibility of bike-sharing enables it to effectively bridge gaps in traditional transit coverage, mitigating limitations in spatial reach and operational frequency (Bruzzone et al., 2021). Particularly at metro and bus transit stations, bike-sharing significantly reduces transfer times, thereby enhancing the connectivity and attractiveness of the overall multimodal public transit network (Liu et al., 2024). The utility of bike-sharing for bridging the "first-mile" and "last-mile" gaps is especially notable in areas where traditional public transit coverage is limited, effectively extending the service radius of metro and bus systems (Huang and Xu, 2023).

In addition to improving transfer efficiency, bike-sharing also plays an essential role in enhancing the equity and accessibility of public transit systems, particularly in peripheral urban areas and communities with limited transportation resources (Duran-Rodas et al., 2021). In neighborhoods at the urban fringe or in areas underserved by conventional transit, bike-sharing provides residents with a flexible and efficient means of reaching the nearest metro or bus station (Guo and He, 2021). This short-distance connectivity service helps to reduce the disparity in transit accessibility across different urban areas, promoting a more equitable distribution of transportation resources and expanding mobility options for underserved populations (Wu and Kim, 2020).

The relationship between bike-sharing and other forms of public transportation, such as buses and metro systems, is characterized by both complementarity and potential competition (Kong et al., 2020). For short-distance trips, the high flexibility

and maneuverability of bike-sharing often make it the preferred mode, especially in scenarios where bus systems have longer wait times or station layouts are less dense (Liu et al., 2022). In these cases, bike-sharing serves as an efficient substitute, providing riders with greater flexibility and reducing their overall travel time (Guo and He, 2020). However, this substitution effect can also impact bus ridership, particularly in areas where bike-sharing is heavily used as a feeder mode to public transit (Godavarthy et al., 2022). To optimize the bike-sharing systems and minimize potential competition, it is necessary to plan for a balanced integration of bike-sharing and public transit (Saltykova et al., 2022). This requires strategic service optimization, ensuring that the complementary relationship between bike-sharing and public transit is maximized while potential conflicts are minimized (Chen et al., 2023).

From a policy perspective, to leverage bike-sharing effectively within multimodal public transit systems, tailored strategies must be implemented in accordance with diverse land-use patterns and transportation service conditions (Luo et al., 2020). For example, in metro or major bus stations, increasing the supply of shared bikes and enhancing the geo-fence of parking can further strengthen transfer efficiency (Cheng et al., 2023). Meanwhile, recognizing the influence of land-use characteristics and identifying target users and potential travel demands are essential for developing adaptive operational strategies (Zhao et al., 2020). These policies ensure that bike-sharing effectively supports multimodal public transit integration.

3. Methodology

3.1 The overall workflow

The primary aim is to understand the complex interplay between the bike-sharing self-loop phenomenon and influential factors such as socioeconomic attributes, land use characteristics, and multimodal public transit conditions across different spatial scales. The analytical framework of this study is as in Fig. 2, which comprises three core components: self-loop intensity calculation and spatial analysis, causal effect estimation, and strategic policy recommendation. Specifically, we first calculate self-loop intensity at multiple spatial scales and employ spatial econometric models to capture the inherent spatial interdependencies within the data. The second component focuses on estimating the causal effects of socioeconomic features, geospatial locations, land use patterns, and multimodal public transit on self-loop by employing double machine learning (DML). The third component pertains to formulating strategic policy recommendations, guided by analysis results derived from the Conditional Average Treatment Effect (CATE), allowing for the exploration of treatment effect heterogeneity by grouping spatial units based on socioeconomic attributes and multimodal public transit conditions.

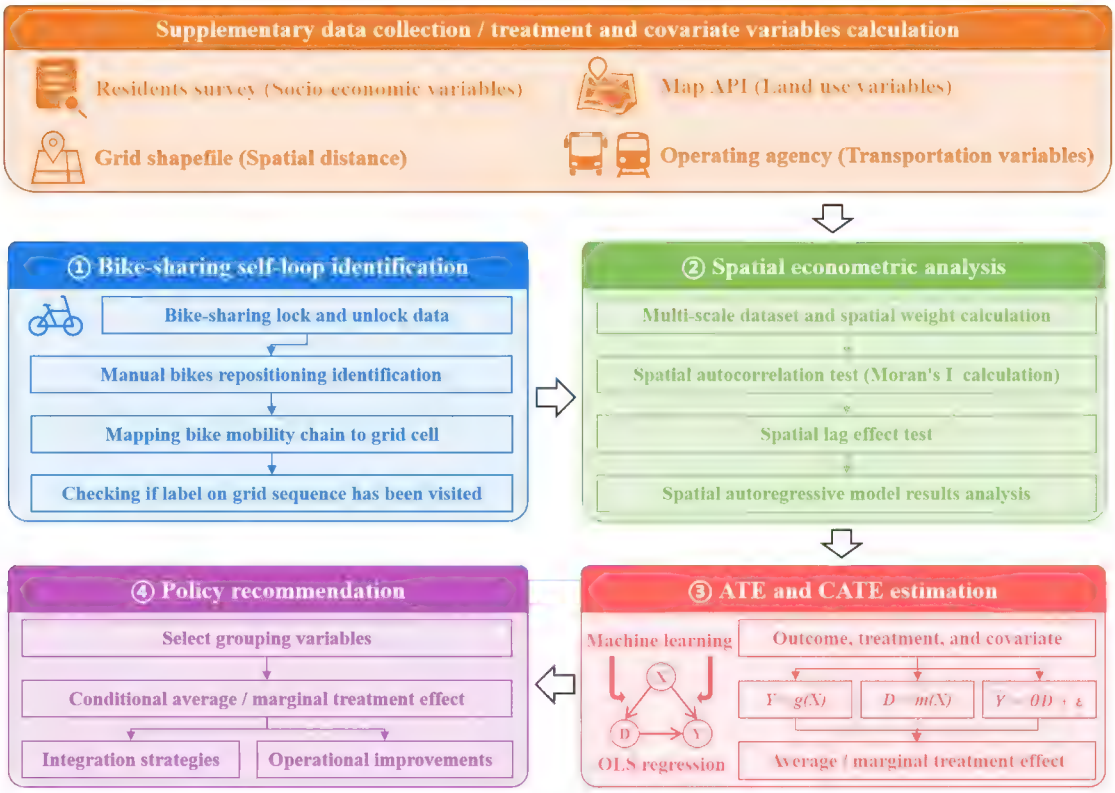

Fig. 2. The research framework of bike-sharing self-loop intensity influencing mechanism.

3.2 Bike-sharing self-loop identification

The bike-sharing self-loop identification begins by preprocessing bike-sharing data. The dataset comprises a series of lock and unlock events for each bike, recorded with the bike ID, timestamp, longitude, and latitude. These records are first sorted by bike ID and the timestamp to establish a bike time-position sequence, enabling the track of each bike's movement. The algorithm proceeds to identify potential bike scheduling instances by service providers. Specifically, if a bike is locked while its recorded location changes significantly, it implies external intervention such as manual bike repositioning. These records are flagged as different bikes in subsequent self-loop identification steps to ensure the accurate reconstruction of bike mobility paths.

Based on the preprocessed bike data, the stationary positions are linked together to create bike mobility chain, providing an aggregated representation of the bike's

movement history across different Origin-Destination (OD) locations. At the same time, the study area is divided into a fine-grained grid. Each stay point is then mapped onto a specific grid cell, generating a series of grid IDs. The self-loop identification is then conducted using a dynamic record dictionary to track each bike's mobility through the grid cells. As the algorithm processes each bike's sequence, it checks whether a grid label has been visited previously. If the grid cell is new, it is added to the dynamic record dictionary with the associated timestamp. If the same grid cell is revisited, the segment between the initial and current visit is recorded as a self-loop. The current timestamp is then updated in the record, ensuring subsequent loops are independently tracked.

This method diverges from other spatial self-loop identification algorithms by requiring only a single pass over the data without missing any self-loop instances, enhancing efficiency and ensuring comprehensive detection (Liu et al., 2018). After identifying all self-loops, the results are aggregated at the station and street scale. Each grid cell is associated with the nearest spatial unit, and the self-loop counts are aggregated accordingly to calculate the metro station and street scale self-loop intensity. The specific pseudocode is shown in Table 2.

Table 2. Pseudocode of self-loop intensity identification.

Algorithm: Bike-sharing Self-loop identification

Input: Bike-sharing lock/unlock records ***T***(bike_id, time, lon, lat); Distance threshold *s*
Output: Identified bike-sharing self-loop records ***C***

```
Initialize empty dictionary C;
Sort T by bike_id and time;
for trip in T do:
    if location change > s while locked then
        Flag as centralized scheduling;
        Continue
    end if
    if consecutive stay within distance s then
        Merge to form stay point;
    else
        Update trip chain;
    end if
```

```
13. end for
14. Map stay points to spatial grid;
15. Initialize dynamic record D;
16. for each grid cell in sequence do
17.     if cell not in D then
18.         Record cell with timestamp;
19.     else
20.         Mark segment as self-loop;
21.     end if
22. end for
return C
```

3.3 Outcome, treatment, and covariate variables

The outcome variable is the bike-sharing self-loop intensity, employed to denote the frequency with which a shared bicycle returns to its original location within a specified time period. To systematically examine the determinants of bike-sharing self-loop behavior, this study identifies two primary categories of treatment variables: land use characteristics and multimodal public transit conditions. Land use characteristics primarily influence cycling demand and trip purposes, whereas multimodal public transit conditions determine the integration of bike-sharing as a first- and last-mile connectivity service to transit networks. Each treatment variable plays a distinct role in shaping self-loop intensity, allowing for a comprehensive analysis of its effects across metro station and street scales.

Specifically, the land use characteristics are divided into two primary groups: location characteristics and Points of Interest (POI) characteristics. Location characteristics include dummy variables indicating whether a spatial unit is located in a Central Urban Area and whether it contains significant functional zones, such as University Areas, Central Business Districts (CBDs), or Transportation Hubs. Central areas typically exhibit elevated bike-sharing demand due to economic concentration and diverse land uses, thereby increasing the frequency of bike-sharing self-loops through repeated, short-distance trips within these densely utilized regions.

To characterize the functional composition of each spatial unit, this study identifies POI attributes measured through the number of Work POIs, Residential POIs, and Commercial/Entertainment POIs. Each POI category provides insight into the types of activities within an area and serves as a factor for potential trip generation. For instance, areas with a higher number of Residential POIs tend to exhibit increased self-loop intensity, as residential areas are associated with regular commuting patterns and frequent, repetitive trips starting and ending at the same location. In contrast, areas with numerous Commercial/Entertainment POIs may exhibit lower self-loop intensity due to the irregular trips related to shopping and leisure activities. Including these POI attributes allows for a more nuanced analysis of how different land use types influence bike-sharing usage patterns.

In order to investigate how multimodal public transit conditions impact bike-sharing self-loop intensity, the metro station ridership and bus station number are also considered as the treatment variable. Metro station ridership serves as a direct indicator of passenger volume, reflecting the integration degree between bike-sharing and urban rail transit. Higher ridership often corresponds with increased demand for first- and last-mile connectivity, leading to greater self-loop intensity. Bus station number measures the bus transit accessibility, indicating the presence of complementary transit infrastructure that supports multimodal public transit connections. It should be noted that metro station ridership is only considered at the metro station scale, while bus station number is included at both the metro station and street scales.

The covariate variables include socioeconomic attributes that are vital for controlling confounding factors in causal analysis. These attributes include the average age of residents, fixed occupations percentage, non-resident population

percentage, and average household vehicle ownership within each spatial unit. The average age is a demographic characteristic that influences bike-sharing usage preference. The fixed occupations percentage serves as an indicator of travel regularity, as individuals with stable employment tend to exhibit consistent commuting patterns, potentially contributing to the self-loop phenomenon. Non-resident population percentage captures specific socioeconomic groups whose mobility behavior may be less routine. Average household vehicle ownership offers an understanding of the availability of private vehicles, which has an inverse relationship with the intensity of bike-sharing self-loop. The definition of outcome, treatment, and covariate variables is shown in Table 3.

Table 3. Definition of outcome, treatment, and covariate variables

Variable	Definition
Outcome variable	
Self-Loop Intensity	Number of occurrences where a shared bike returns to its original grid within a defined time period.
Covariate variables	
Average age	The average age of the population within the spatial unit.
Fixed occupations percentage	Percentage of individuals with stable employment within the spatial unit.
Non-resident population percentage	Percentage of non-resident individuals within the spatial unit.
Average household car ownership	Average number of vehicles owned per household within the spatial unit.
Land use variables	
Central area dummy	The indicator variable is whether the spatial unit is located in the central urban area.
University area dummy	The indicator variable is whether the spatial unit contains a university.
Central business district area dummy	The indicator variable is whether the spatial unit contains a central business district (CBD).
Transportation hub area dummy	Indicator variable for whether the spatial unit contains a major transportation hub.
Work POI intensity	Number of work-related POI within the spatial unit.
Residential POI intensity	Number of residential POI within the spatial unit.
Commercial POI intensity	Number of commercial-related POI within the spatial unit.
Metro station ridership	Average daily passenger ridership of metro stations (Metro scale only).
Bus station number	Number of bus stops within the spatial unit.

3.4 Spatial econometric analysis

The Moran's I test is employed to examine whether bike-sharing self-loop intensity exhibits spatial clustering across different urban spatial units. The Global

Moran's I statistic is calculated to assess the spatial autocorrelation at the metro station and street scale. This test helps identify if bike-sharing self-loop intensity shows a significant spatial clustering phenomenon (high-intensity areas tend to be surrounded by other high-intensity areas). The formulation of Moran's I is as follows:

$$I = \frac{n}{\sum_i \sum_j w_{i,j}} \times \frac{\sum_i \sum_j w_{i,j}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2} \tag{1}$$

where y_i represents the bike-sharing self-loop intensity for unit i, $\bar{y}$ is the average self-loop intensity, and $w_{i,j}$ is the spatial weight between units i and j. A significantly positive Moran's Index highlights the necessity to address spatial dependencies in subsequent modeling.

The Spatial Autoregressive (SAR) model is used to account for spatial lag effects, incorporating a spatially lagged dependent variable to capture the influence of neighboring regions on the bike-sharing self-loop intensity of a given spatial unit. The SAR model is formulated as:

$$y = \rho W y + X\beta + \varepsilon \tag{2}$$

where y denotes the bike-sharing self-loop intensity, Wy represents the spatially lagged dependent variable, ρ is the spatial autoregressive coefficient, X is the matrix of covariates, β is the vector of coefficients, and ε is the error term. The term ρWy captures the spillover effect from neighboring spatial units, reflecting the spatial dynamics of bike-sharing usage.

3.5 Treatment effects estimation and policy recommendation

The double machine learning approach(DML) is employed to estimate the causal effects (Bach et al., 2024). The DML framework provides unbiased causal estimates and serves as a foundation for effective policy recommendations. The relationship

between outcome, treatment, and covariates can be formulated as Eq. (2).

$$Y = g(X) + \theta D + \varepsilon \tag{3}$$

where Y represents the outcome variable (bike-sharing self-loop intensity), D is the treatment variable, $g(X)$ represents the effect of covariates, θ represents the causal effect parameters to be estimated, ε is the error term. Similarly, the treatment assignment model is defined as Eq. (3).

$$D = m(X) + \upsilon \tag{4}$$

where $m(X)$ represents the propensity score or the conditional expectation of the treatment given covariates, υ is the error term for the treatment model. After the functions $\hat{g}(X)$ and $\hat{m}(X)$ are estimated, the residuals for both the outcome and the treatment are computed as Eq. (4)-(5).

$$\hat{Y} = Y - \hat{g}(X) \tag{5}$$

$$\hat{D} = D - \hat{m}(X) \tag{6}$$

The residuals represent the components of the outcome and treatment which are uncorrelated with the covariates, thus reducing the bias introduced by confounding factors. The treatment effect θ is then estimated using the orthogonalized residuals through the following formula:

$$\theta = \frac{\sum_{i=1}^{n} \hat{D}_i \hat{Y}_i}{\sum_{i=1}^{n} \hat{D}_i^2} \tag{7}$$

where $\hat{D}_i$ and $\hat{Y}_i$ are the residuals for each observation i, n is the number of observations. This formulation ensures that the estimated treatment effect is robust to the influence of confounders, providing an unbiased estimation of $\hat{\theta}$ through the use of orthogonalized residuals. The baseline estimators in the DML framework in this

study is Xgboost.

In the DML framework, the explanation of the Average Treatment Effect (ATE) and Conditional Average Treatment Effect (CATE) varies depending on whether the treatment variable D is discrete or continuous. For a discrete treatment, where $D=1$ indicates treated and $D=0$ indicates untreated, the ATE is defined as:

$$ATE = \mathrm{E}\left[Y(1)\right] - \mathrm{E}\left[Y(0)\right] \tag{8}$$

The ATE provides an estimate of how bike-sharing self-loop intensity changes if a discrete intervention policy is applied across all units. For continuous treatments, the ATE generalizes to the Marginal Treatment Effect (MTE), defined as:

$$MTE = \frac{\mathrm{E}\left[Y(D)\right]}{\partial D} \tag{9}$$

The MTE captures the change rate in the outcome given a marginal increase in the treatment level, providing an understanding of how changes in land use or transportation conditions may affect the bike-sharing self-loop intensity (Zhou and Xie, 2020). The CATE further extends these concepts by conditioning on covariates X . For discrete treatment, the CATE is expressed as:

$$CATE(X) = \mathrm{E}\left[Y(1) - Y(0) \mid X\right] \tag{10}$$

This measures the treatment effect for subpopulations defined by specific covariates (socioeconomic attributes). For continuous treatments, CATE is defined as:

$$CATE(X) = \frac{\mathrm{E}\left[Y(D) \mid X\right]}{\partial D} \tag{11}$$

It provides the marginal effect of the treatment on the outcome for specific subgroups, allowing us to identify the heterogeneous impacts of interventions.

4. Result analysis

4.1 Study area and descriptive analysis

Shanghai, China, was selected as the study case to analyze the mechanism of bike-sharing self-loop at metro stations and street scales. The analysis covers 407 metro stations and 229 street segments, providing an understanding of bike-sharing patterns across different urban environments. The primary dataset includes bike-sharing trip data collected from August 1 to August 14, 2023, representing two weeks of unlocking and locking activities for 392,083 bicycles, comprising a total of 52, 897, 906 records. The self-loop identification and calculation periods were set to one week, with daily self-loop intensity calculated as the outcome variable, and all manual bike reallocations have already been excluded to ensure accuracy. The spatial matching relationship among the grid, metro station, and street segment is shown as Fig 3.

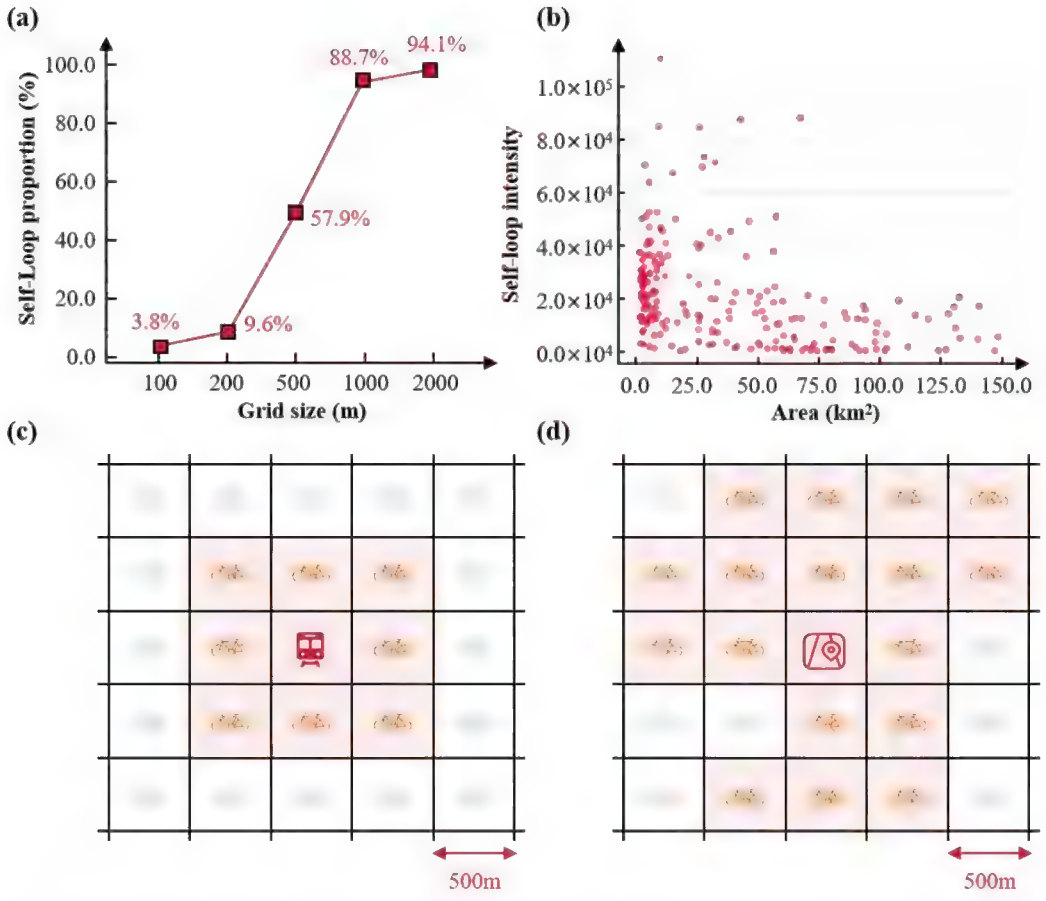

Fig. 3. Spatial matching relationship among grid, metro station, and street segment. **(a)** Relationship between grid size and bike-sharing self-loop proportion. **(b)** Relationship between street segment area and self-loop intensity. **(c)** Matching relationship between

grid and metro station. **(d)** Matching relationship between grid and street segment.

Fig. 3(a) illustrates the relationship between grid size and the bike-sharing self-loop proportion. The bike-sharing self-loop proportion is defined as the ratio of the sum self-loop number across the grids to the total number of bike-sharing trips. Grid size plays a crucial role in the identification of self-loops: when the grid size is too small, it becomes difficult to identify the return trips of shared bikes, whereas when the grid size is too large, the majority of trips fall within self-loops, rendering the analysis meaningless. Therefore, we have selected a moderate grid size of 500 meters.

The identified grid scale self-loops are further aggregated to derive self-loop intensity at the metro station and street segment scales. The spatial matching relationships at these two scales are illustrated in Fig. 3(c) and 3(d), respectively. At the metro station scale, the self-loop intensity is calculated by aggregating the self-loop intensity of the nine surrounding grids centered on a particular metro station. At the street scale, the self-loop intensity is aggregated to the street segments according to the spatial matching relationship between the grids and the street segment. To analyze the impact of spatial unit size on self-loop intensity, we plotted a scatter diagram depicting the relationship between area and self-loop intensity at the street scale, as shown in Fig. 3(b). The scatter plot reveals no significant correlation between area and self-loop intensity, suggesting that the heterogeneity in self-loop intensity across different spatial units is attributable to intrinsic mechanisms rather than variations in area.

The spatial distribution of bike-sharing trip and self-loop intensity exhibits significant heterogeneity across different scales, as shown in Figs. 4 and 5. At the grid scale, trip intensity is primarily concentrated in the southwestern part of the city and

peripheral areas, while self-loop intensity is centered in the urban core. At the metro station scale, high self-loop intensity is also concentrated in the urban core. However, at the street segment scale, elevated self-loop intensity is observed both in the central areas and around newly developed towns. It is noteworthy that metro station such as Jiading New Town, Minhang Jiaotong University and Xinzhuang exhibit trends that deviate from the overall pattern, which may be attributed to the localized concentration of bike deployment and do not affect the overall analysis.

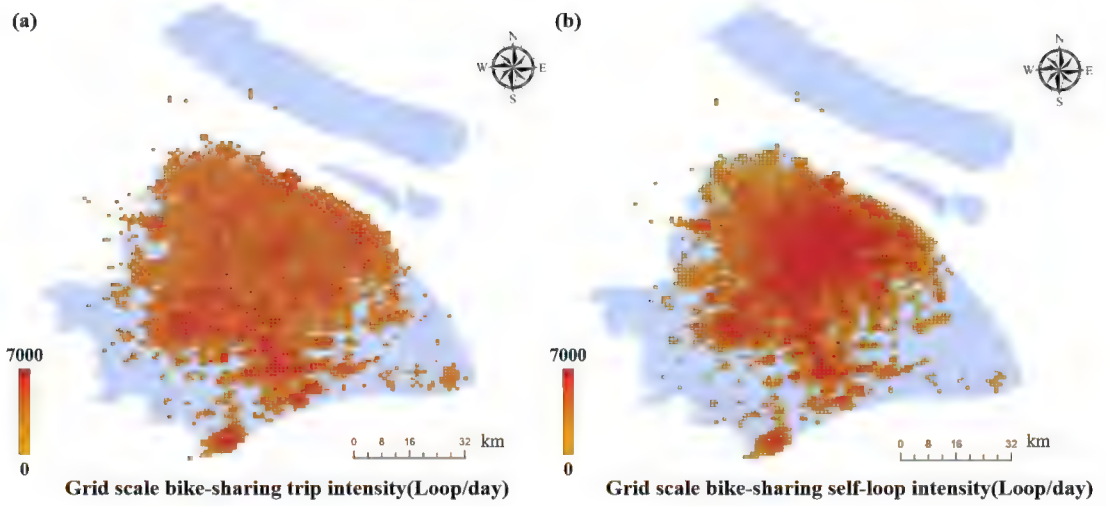

Fig.4. Distribution of bike-sharing trip and self-loop intensity at the grid scales. **(a)** Bike-sharing trip intensity. **(b)** Bike-sharing self-loop intensity.

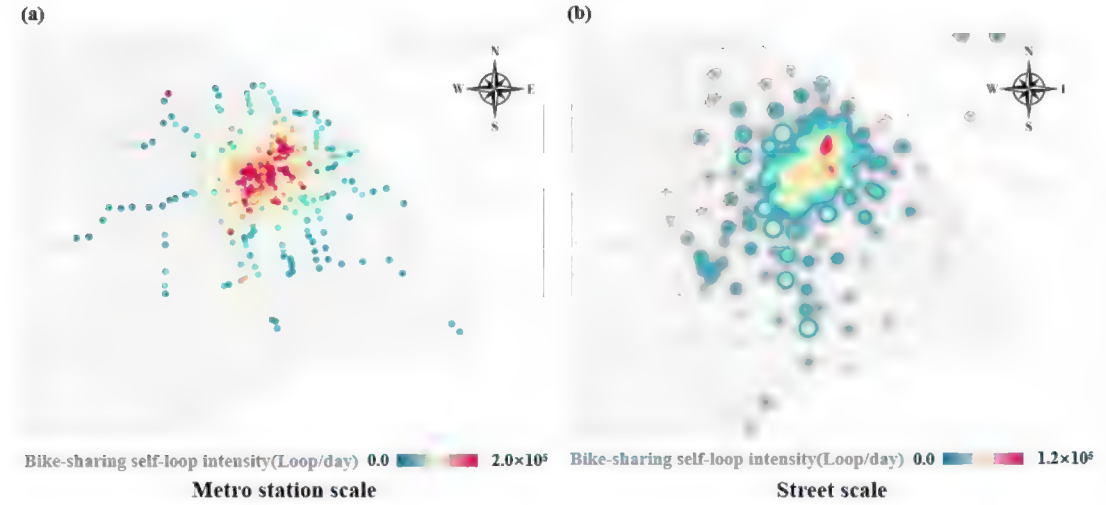

Fig.5. Distribution of bike-sharing self-loop intensity at metro station and street scales. (a) Metro station scale. (b) Street scale.

Land use attributes were collected using a map API service to acquire detailed

Points of Interest (POI) and information on major functional areas across Shanghai, reflecting the land use diversity within each spatial unit. Metro station ridership data and bus stop locations were obtained for the same period to provide an understanding of multimodal public transit conditions that potentially influence bike-sharing behavior. Socioeconomic attributes were obtained from the Shanghai Residents Survey, providing a detailed representation of the socio-demographic factors. The distributions of socioeconomic attributes at metro station and street scales are shown in Fig. 6, while the descriptive statistics of each spatial scale are provided in Tables 4 and 5.

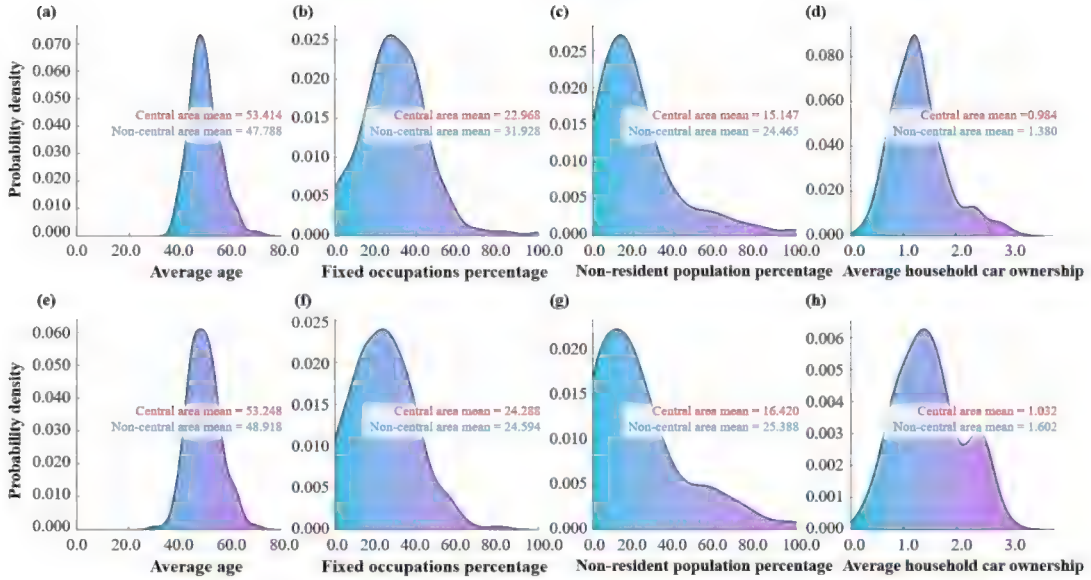

Fig.6. The distributions of socioeconomic attributes at metro station and street scales. **(a)** Average age at metro station scale. **(b)** Fixed occupations percentage at metro station scale. **(c)** Non-resident population percentage at metro station scale. **(d)** Average household car ownership at metro station scale. **(e)** Average age at street scale. **(f)** Fixed occupations percentage at street scale. **(g)** Non-resident population percentage at street scale. **(h)** Average household car ownership at street scale.

Table 4. Descriptive statistics of variables at the metro station scale.

Variables	Unit	Mean	Std	Min	Max
Bike-sharing Self-Loop intensity	loop/day	2756.118	3279.035	5.286	19572.571
Average age	num	48.991	5.780	37.000	70.000
Fixed occupations percentage	%	30.013	15.921	0.000	100.000
Non-resident population percentage	%	22.473	20.083	0.000	100.000
Average household car ownership	vehicle	1.295	0.532	0.172	3.200
Work POI intensity	num	80.855	89.110	0.000	565.000

Residential POI intensity	num	54.885	58.523	0.000	317.000
Commercial POI intensity	num	278.690	321.037	0.000	2062.000
Central area dummy	-	0.214	0.410	0.000	1.000
University area dummy	-	0.300	0.459	0.000	1.000
Central business district area dummy	-	0.388	0.488	0.000	1.000
Transportation hub area dummy	-	0.140	0.347	0.000	1.000
Bus station number	station	4.637	4.605	0.000	23.000
Metro station ridership	people/day	29471.668	31866.844	550.000	247111.857

Table 5. Descriptive statistics of variables at the street scale.

Variables	**Unit**	**Mean**	**Std**	**Min**	**Max**
Bike-sharing Self-Loop intensity	loop/day	2290.359	2156.268	1.429	11651.571
Average age	num	49.750	6.172	30.942	70.000
Fixed occupations percentage	%	24.535	15.129	0.000	80.000
Non-resident population percentage	%	23.665	22.370	0.000	100.000
Average household car ownership	vehicle	1.493	0.623	0.241	3.018
Work POI intensity	num	571.803	502.005	11.000	3148.000
Residential POI intensity	num	252.092	181.186	5.000	1620.000
Commercial POI intensity	num	1391.533	1036.171	7.000	5172.000
Central area dummy	-	0.192	0.395	0.000	1.000
University area dummy	-	0.354	0.479	0.000	1.000
Central business district area dummy	-	0.384	0.487	0.000	1.000
Transportation hub area dummy	-	0.153	0.361	0.000	1.000
Bus station number	station	22.041	18.948	0.000	83.000

Although the street scale covers a larger geographic area compared to the metro station scale, metro stations are more concentrated in central urban areas, leading to a higher average bike-sharing self-loop intensity. For socioeconomic attributes, the metro station scale is characterized by a slightly younger population compared to the street scale, as evidenced by a lower average age. The non-resident population percentage at metro stations is also marginally lower, while the fixed occupations percentage is significantly higher—5.478% greater than at the street scale. These trends indicate that metro station areas are more likely to encompass the workplace, thus attracting a stable working population in regular commuting.

Regarding land use characteristics, the broader spatial coverage of the street scale naturally leads to a higher absolute number of Points of Interest (POI) compared to

the metro station scale. However, the composition of POIs varies markedly between the two scales. Commercial and entertainment POIs are notably concentrated around metro stations, with a ratio of 3.475 commercial POIs per work-related POI at the metro station scale, compared to a ratio of 2.436 at the street scale. For location variables, metro stations are more likely to be located in central urban areas compared to street scale. The proportion of metro stations located near universities, CBDs, or transportation hubs is similar to that of street scale. For transportation conditions, the average daily ridership at metro stations is nearly 30,000 passengers. Conversely, the bus station's number is greater at the street scale due to broader geographic coverage.

4.2 Spatial dynamics and land use influencing mechanisms on self-loop intensity

The SAR model results reveal distinct spatial dynamics between metro station and street scales, highlighting how spatial dependencies, socioeconomic attributes, land use characteristics, and multimodal public transit conditions shape bike-sharing patterns, as shown in Tables 6 and 7.

Table 6. Spatial autoregressive results of bike-sharing self-loop intensity at metro station scale.

Variable	Coef.	Std. Err.	z	P>|z|
Socioeconomic variables				
Average age	0.037 ***	0.012	3.150	0.002
Fixed occupations percentage	0.005	0.003	1.470	0.142
Non-resident population percentage	0.001	0.003	0.350	0.730
Log(Average household car ownership)	-0.528 ***	0.119	-4.430	0.000
Land use and geospatial location variables				
Log(Work POI intensity)	0.030	0.060	0.490	0.624
Log(Residential POI intensity)	0.550 ***	0.086	6.400	0.000
Log(Commercial POI intensity)	-0.013	0.060	-0.220	0.827
Central area dummy	-0.005	0.153	-0.030	0.973
University area dummy	0.315 ***	0.107	2.950	0.003
Central business district area dummy	0.101	0.101	1.000	0.318
Transportation hub area dummy	0.354 **	0.145	2.440	0.015
Transportation variables				

Log(Bus station number)	0.179 ***	0.067	2.670	0.008
Log(Metro station ridership)	0.242 ***	0.055	4.400	0.000
Constant	1.007	0.851	1.180	0.236
Spatial lag term	0.013	0.085	0.150	0.882
Observation			407	
Pseudo R2			0.683	

Table 7. Spatial autoregressive results of bike-sharing self-loop intensity at street scale.

Variable	**Coef.**	**Std. Err.**	**z**	**P>|z|**
Socioeconomic variables				
Average age	-0.007	0.018	-0.380	0.701
Fixed occupations percentage	0.000	0.006	0.070	0.943
Non-resident population percentage	-0.002	0.004	-0.420	0.678
Log(Average household car ownership)	-0.235	0.198	-1.190	0.235
Land use and geospatial location variables				
Log(Work POI intensity)	0.038	0.126	0.300	0.766
Log(Residential POI intensity)	1.005 ***	0.192	5.230	0.000
Log(Commercial POI intensity)	-0.552 ***	0.169	-3.260	0.001
Central area dummy	-1.028 ***	0.284	-3.620	0.000
University area dummy	0.109	0.159	0.690	0.493
Central business district area dummy	0.324 **	0.165	1.960	0.050
Transportation hub area dummy	0.137	0.206	0.660	0.507
Transportation variables				
Log(Bus station number)	0.077	0.076	1.020	0.308
Constant	4.274 ***	1.317	3.250	0.001
Spatial lag term	0.240 ***	0.052	4.620	0.000
Observation			229	
Pseudo R2			0.508	

At the metro station scale, the spatial lag term was found to be insignificant, probably due to the physical distance between metro stations. In contrast, at the street scale, the spatial lag term was significant, with a coefficient of 0.240, indicating spatial spillover effects where bike-sharing self-loop intensity in one street segment has correlation with neighboring street segments.

Furthermore, the SAR results reveal differences in significance levels between the two spatial scales. At the metro station scale, socioeconomic, land use, and transportation factors all exerted significant effects on self-loop intensity, whereas, at the street scale, land use characteristics emerged as the most critical factor. Land use

influences bike-sharing activity across all spatial scales, while socioeconomic and transportation factors contribute more specifically to metro station-scale variations.

Focusing on the Points of Interest (POI) attributes, the results show consistent coefficient directions across both spatial scales but reveal differing levels of significance and elasticity. At the street scale, a 1% increase in Residential POI intensity corresponds to a 1.005% increase in self-loop intensity, while a 1% increase in Commercial POI intensity leads to a 0.552% decrease in self-loop intensity. These findings suggest that residential areas encourage more regular bike-sharing usage, likely driven by commuting or first- and last-mile connectivity, whereas commercial areas lead to more irregular usage patterns, such as for leisure or shopping trips. The influence of POI attributes at the metro station scale follows the same directional pattern but with lower elasticity and significance, implying a weaker influence compared to the broader street scale.

For location-specific attributes, significant effects were found at both metro station and street scales, though the variables differed. At the metro station scale, the University area dummy and transportation hub area dummy were significant, resulting in 0.315% and 0.354% increases in bike-sharing self-loop intensity, respectively. Conversely, at the street scale, the Central area dummy and CBD area dummy showed significant effects. The Central area dummy had a negative effect on self-loop intensity, indicating that bike-sharing self-loop activity is more prevalent in suburban areas, while the CBD area dummy had a positive effect, resulting in a 0.324% increase in self-loop intensity. These results reflect the heterogeneity between scales, with metro station areas influenced by proximity to specific facilities, while macro-level land use characteristics dominate at the street scale.

Multimodal public transit connectivity also plays a crucial role, particularly at the metro station scale. A 1% increase in bus station numbers and metro station ridership leads to 0.242% and 0.179% increases in self-loop intensity, respectively, emphasizing the importance of bike-sharing in enhancing first- and last-mile connectivity around metro stations. However, at the street scale, the influence of bus station numbers was not significant, suggesting that while bus services are vital for localized station-level connectivity, their impact becomes negligible at broader spatial scales dominated by land use and the built environment.

The results highlight the nuanced and scale-dependent influences on bike-sharing self-loop intensity. The findings underline how different spatial factors exert variable influences depending on the scale of analysis, with land use being a major determinant at both scales, while socio-economic and transportation factors play a more specific role at the metro station level.

4.3 Scale-specific determinants: residential POI intensity and public transit as key drivers of bike-sharing self-loop intensity

The causal analysis of bike-sharing self-loop intensity was conducted using a double machine learning (DML) framework, taking bike-sharing self-loop intensity as the outcome variable, socioeconomic attributes as covariates, and land use characteristics along with transportation conditions as the treatment variables. This approach enabled a comprehensive estimation of the Average Treatment Effect (ATE) for categorical variables and the Marginal Treatment Effect (MTE) for continuous treatment variables, providing robust insights into influencing mechanisms. Results from the DML estimations for both metro station and street scales are presented in Tables 8 and 9. The R^2 values of 0.806 for the metro station scale and 0.699 for the street scale exceed those from the SAR model (0.683 and 0.508, respectively),

indicating a superior fit and enhanced explanatory power compared to traditional spatial econometric approaches.

Table 8. Double machine learning estimation results of bike-sharing self-loop intensity at metro station scale (Utilizing socioeconomic attributes as covariate variables).

Treatment variable	Coef.	Std. Err.	t	P>|t|
Land use variables				
Log(Work POI intensity)	0.052	0.022	2.385	0.017
Log(Residential POI intensity)	0.460 ***	0.041	11.344	0.000
Log(Commercial POI intensity)	-0.049	0.024	-2.037	0.042
Central area dummy	0.167 ***	0.056	3.000	0.003
University area dummy	0.303 ***	0.033	9.223	0.000
Central business district area dummy	0.100 ***	0.034	2.924	0.004
Transportation hub area dummy	0.242 ***	0.059	4.095	0.000
Transportation variables				
Log(Bus station number)	0.450 ***	0.046	9.692	0.000
Log(Metro station ridership)	0.600 ***	0.063	9.531	0.000
Observations		407		
R^2		0.806		
Adj R^2		0.806		

Table 9. Double machine learning estimation results of bike-sharing self-loop intensity at street scale (Utilizing socioeconomic attributes as covariate variables).

Treatment variable	Coef.	Std. Err.	t	P>|t|
Land use variables				
Log(Work POI intensity)	0.005	0.051	0.105	0.916
Log(Residential POI intensity)	1.055 ***	0.151	6.997	0.000
Log(Commercial POI intensity)	-0.046	0.086	-0.537	0.591
Central area dummy	-0.071	0.081	-0.886	0.376
University area dummy	0.170 ***	0.048	3.579	0.000
Central business district area dummy	0.265 ***	0.043	6.094	0.000
Transportation hub area dummy	0.001	0.103	0.005	0.996
Transportation variables				
Log(Bus station number)	0.489 ***	0.105	4.664	0.000
Observations		229		
R^2		0.699		
Adj R^2		0.698		

As illustrated in Tables 8 and 9, the coefficient directions were consistent across both scales, though differences in significance and treatment effect values were evident. For POI attributes, Residential POI intensity showed significant positive effects on self-loop intensity at both scales, with more pronounced marginal effects at

the street scale. A 1% increase in Residential POI intensity was linked to a 0.460% increase in self-loop intensity at the metro station scale and a 1.055% increase at the street scale. These findings emphasize that residential land use contributes substantially to bike-sharing activity, particularly at the street scale, highlighting the importance of residential connectivity.

For location attributes, divergent patterns emerged between the two spatial scales. At the metro station scale, all four location-related dummies exhibited significant positive effects, suggesting a widespread impact of these features on self-loop intensity. Specifically, the Central area dummy had an ATE of 0.167%, while the University area dummy (0.303%), Transportation hub area dummy (0.242%), and CBD area dummy (0.100%) also demonstrated significant effects. These findings indicate that universities and transportation hubs are key determinants of bike-sharing self-loops at the station level. Conversely, at the street scale, only the University area and CBD area dummies remained significant, with ATEs of 0.170% and 0.265%, respectively, implying that, at a broader scale, CBDs act as major land-use drivers of bike-sharing self-loop intensity, while other localized attributes lose their prominence.

Multimodal public transit connectivity emerged as a crucial determinant at both spatial scales, with significant positive MTEs observed. A 1% increase in bus station number resulted in a 0.450% increase in self-loop intensity at the metro station scale and a 0.489% increase at the street scale. Additionally, a 1% increase in metro station ridership led to a 0.600% increase in self-loop intensity at the metro station scale. These results underscore the pivotal role of public transit connectivity in shaping bike-sharing behavior, reinforcing the notion that improving bus and metro accessibility can significantly enhance bike-sharing self-loop intensity.

5. Discussion

5.1 Bike-public transit integration strategies at metro station scale

The bike-sharing self-loop intensity could serve as an effective metric to assess integration between bike-sharing and multimodal public transit systems, reflecting the efficiency of first- and last-mile connectivity. To examine how land use patterns influence bike-sharing integration across different public transit service conditions, metro station ridership and bus station numbers were used as grouping criteria. Fig. 7 presents the treatment effects of land use variables across these groups, and the spatial unit distribution at metro scale is shown in Fig. 8.

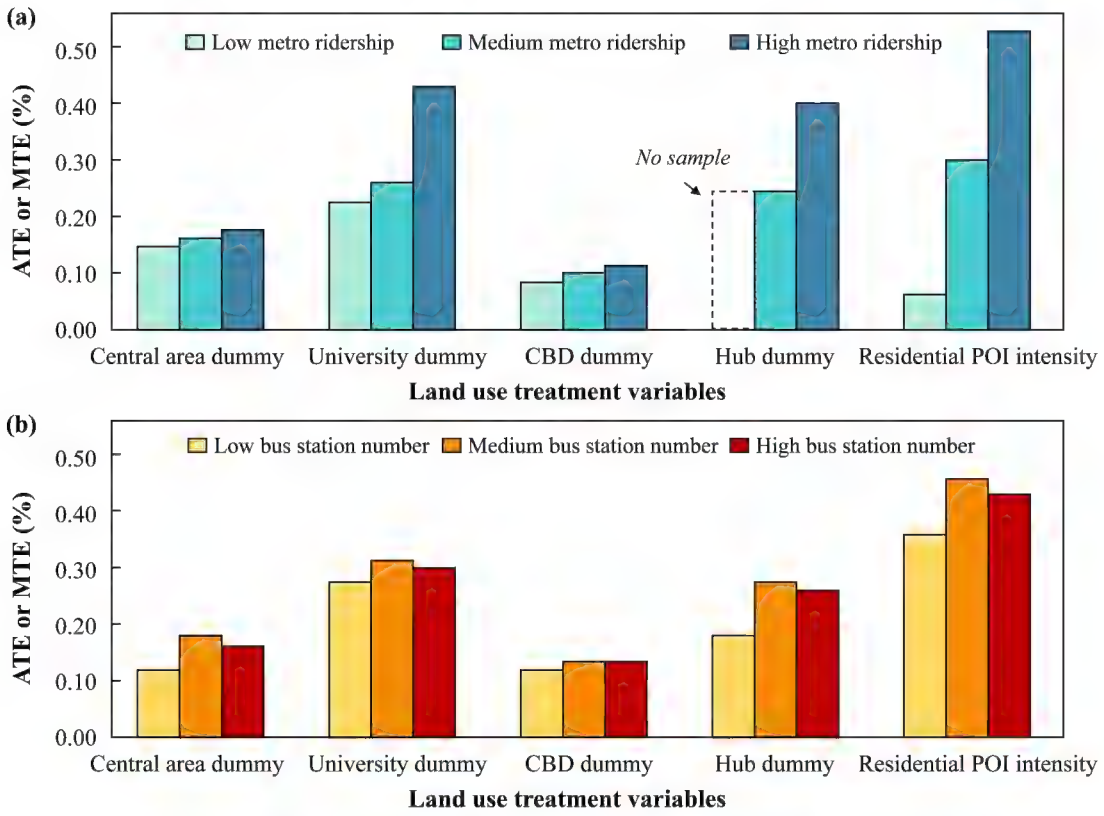

Fig. 7. Variation in ATE/MTE of land use variables across different transportation service conditions at the metro station scale: (a) Grouped by metro station ridership (b) Grouped by bus station number.

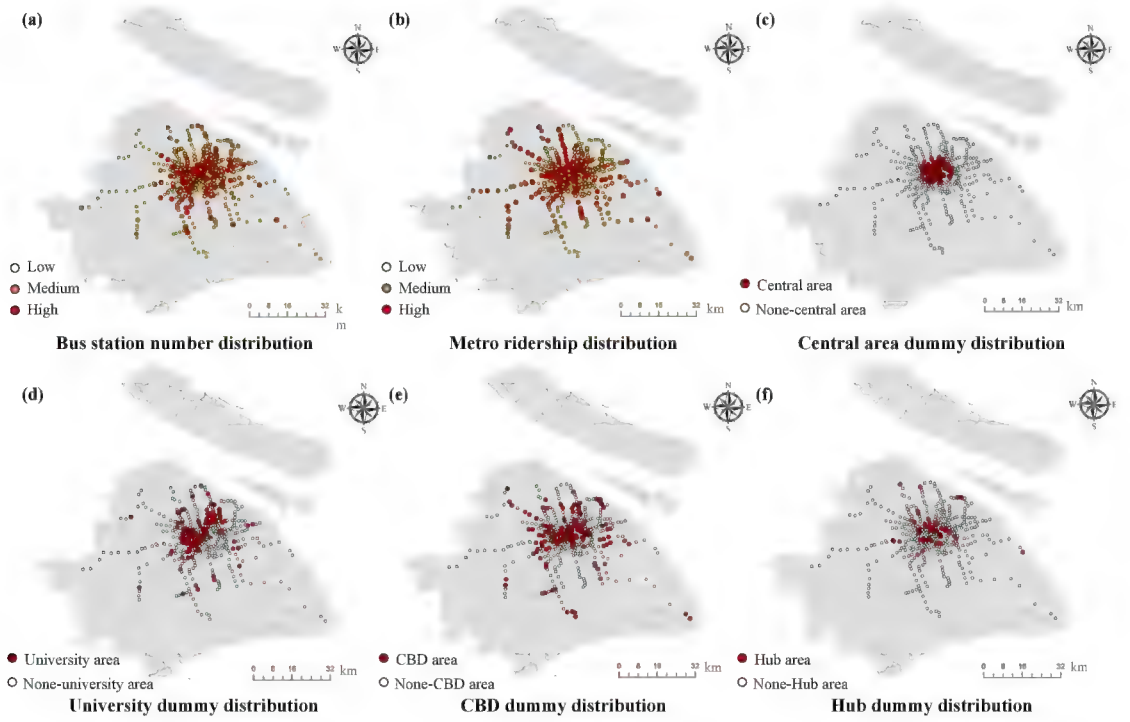

Fig. 8. Spatial unit distribution at metro scale **(a)** Bus station number as the group variable. **(b)** Metro ridership as the group variable. **(c)** Central area dummy as the group variable. **(d)** University dummy as the group variable. **(e)** CBD dummy as the group variable. **(f)** Hub dummy as the group variable.

The analysis shows that stations in the high metro ridership group generally demonstrate greater ATE (for dummy variables) and MTE (for continuous variables) compared to the medium and low ridership groups. However, the extent of inter-group differences varies by land use variable. Specifically, Central area and CBD dummies show low ATE values with minimal variability across groups, suggesting a limited but steady effect on bike-sharing intensity. In contrast, the University dummy, the Transportation hub dummy, and the Residential POI intensity reveal substantial treatment effects and notable inter-group differences, particularly around high metro ridership stations.

These findings lead to targeted strategies for improving bike-sharing integration at high-ridership metro stations. For stations near universities, such as Fudan University and Shanghai Jiao Tong University, increasing bike availability during

peak times can ease congestion and enhance first- and last-mile connections (Caggiani et al., 2020). Collaborations between bike-sharing providers and universities to offer subsidies specifically for bike-to-metro transfers could also enhance convenience and boost connectivity around university areas (Gao et al., 2021).

In transportation hub areas like Shanghai Hongqiao Railway Station and Pudong International Airport, enhancing wayfinding and dedicated bike lanes can improve connectivity, while flexible geo-fencing for bike-sharing parking can reduce spatial redistribution and rebalancing needs (Zhuang et al., 2021). These hubs are major points of multimodal transfer, and the addition of clear wayfinding for cyclists could ensure smoother transitions between different transport modes .

For high-density residential areas around high-ridership metro stations, such as those near Zhangjiang High-Tech Park and Baoshan Wanda, adjusting bike redistribution schedules to match commuting peaks and establishing dedicated cycling lanes connecting residential zones to metro stations can significantly enhance convenience and connectivity (Zhang et al., 2023). These residential clusters require careful management of bike supply to meet peak hour demands, thereby ensuring that residents can conveniently connect to metro services.

Fig. 7(b) reveals that while ATE/MTE for each land use variable are similar to those in Fig. 7(a), the ATE/MTE at varying bus station number groups exhibit distinct differences across land use variables. Specifically, the low bus station number group shows the lowest ATE and MTE, whereas the medium bus station number group exhibits a significant increase in treatment effects for the University area dummy, Transportation hub dummy, and Residential POI intensity. For the high bus station number group, the ATE and MTE effects decrease slightly compared to the medium group. This indicates that in areas like Jiading New Town and Songjiang New City,

where bus connectivity is limited, bike-sharing can serve as a vital supplement, effectively filling gaps left by bus transit services. The relationship between bus stations and bike-sharing is characterized by both competition and complementarity—where fewer bus stations lead to bike-sharing serving as a supplement, more bus stations result in a competitive relationship between the two modes. This phenomenon is particularly evident in areas characterized by high Residential POI intensity.

It is recommended to avoid expanding the shared bike resources where bus stations already provide extensive coverage. Instead, resources should be concentrated in areas with high metro ridership but fewer connecting bus stations (Guo and He, 2020). In these locations, bike-sharing can serve as an essential first- and last-mile connector, effectively compensating for limited bus access and improving overall system efficiency. Partnerships between transit authorities and bike-sharing providers could also offer combined fare incentives or integrated travel passes, encouraging users to adopt bike-sharing where traditional bus services are insufficient (Jin et al., 2022a).

5.2 Bike-sharing operational improvements at street scale

The empirical analysis at the street scale demonstrates a significant positive treatment effect of residential POI intensity on bike-sharing self-loop intensity. It suggests that residential land use patterns are a major determinant of bike-sharing activities. However, given the diversity of socioeconomic attributes among residential units, a more nuanced examination is required to understand how the MTE of residential POI intensity varies across socioeconomic groups, as illustrated in Fig. 9 and 10. For age distribution, streets in the western regions are predominantly composed of middle-aged residents, while those in the eastern regions are primarily

composed of older populations. Regarding car ownership, streets in the central urban area and Pudong New Area exhibit higher average household car ownership. The distribution of streets grouped by non-resident population percentage or fixed occupation percentage does not show significant spatial clustering, which may be more influenced by land use type.

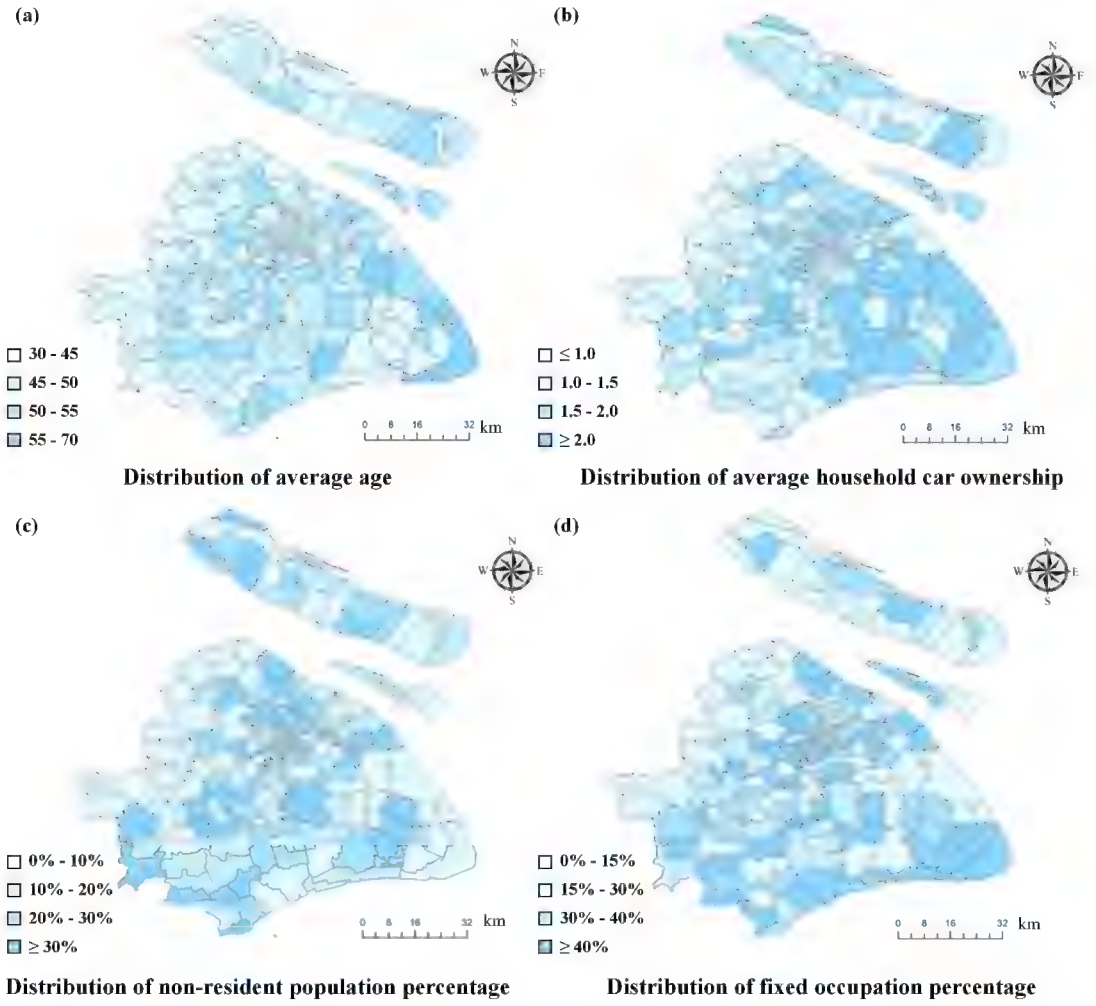

Fig. 9. Spatial distribution of street segments grouped by socioeconomic attributes. **(a)** Distribution of streets grouped by average age. **(b)** Distribution of streets grouped by average household car ownership. **(c)** Distribution of streets grouped by non-resident population percentage. **(d)** Distribution of streets grouped by fixed occupation percentage.

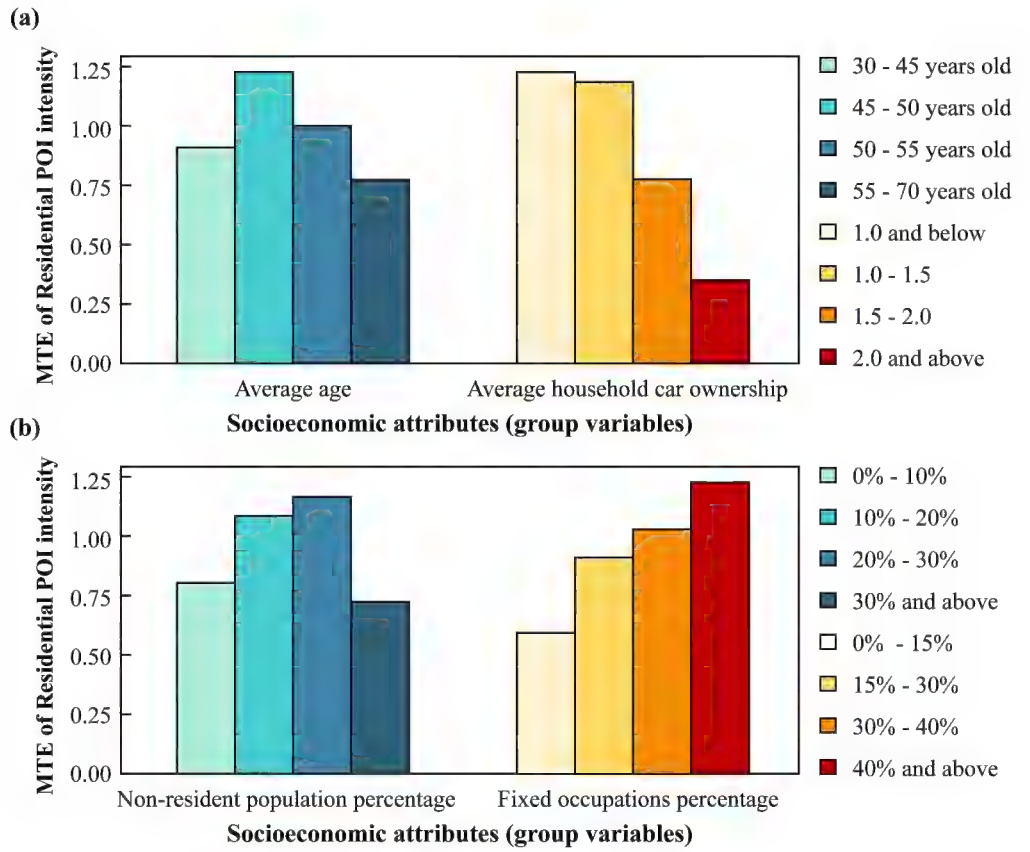

Fig. 10. Variations in residential POI Intensity MTE among different socioeconomic attribute groups at the street scale. **(a)** Average age and average household car ownership. **(b)** Non-resident population percentage and fixed occupation percentage.

Fig. 10 reveals notable variations in the MTE of residential POI intensity across different socioeconomic groups. Streets with an average age of 45-50 years old demonstrate the highest MTE, likely due to their commuting behaviors. Conversely, areas with predominantly older residents (55-70 years) exhibit the lowest treatment effects, probably due to the physical limitations. The analysis also shows that streets with lower car ownership (1.0 or below) have the highest MTE values, while higher car ownership correlates with reduced reliance on bike-sharing. Streets with moderate non-resident populations (20-30%) and those with a high percentage of fixed occupations (over 40%) also show substantial MTE, highlighting the importance of targeted bike-sharing operation strategies to accommodate transient populations.

To optimize bike-sharing at the street scale, several policy recommendations are proposed. In areas dominated by middle-aged residents (45-50 years old), expanding

bike-sharing infrastructure is essential. Ensuring adequate bike availability during commuting hours would facilitate broader adoption. For low car ownership areas, integrating bike-sharing with other transit options such as regular buses or customized shuttles can maximize the bike-sharing service usage rate (Bieliński et al., 2021). Improved cycling infrastructure, including dedicated bike lanes connecting residential zones to essential amenities, can ensure that bike-sharing becomes a viable short-distance travel alternative to personal vehicles.

In streets with a moderate non-resident population percentage, it is necessary to maintain flexibility in bike-sharing operations. Real-time bike-sharing allocation should be employed to ensure availability across different times of the day, addressing fluctuating demand effectively (Jin et al., 2022b). Additionally, in fixed occupation-dense areas, integrating bike-sharing into commuting routines by establishing workplace-adjacent parking geo-fence can enhance bike-sharing adoption. Providing discounted bike-sharing memberships would encourage more commuters to consider bike-sharing as part of their daily routine.

6. Conclusion

This study conducts a multiscale analysis of the bike-sharing self-loop phenomenon, investigating its influencing mechanisms at both metro station and street scales. We used a spatial autoregressive model and a double machine learning framework to capture the complex interactions between socioeconomic attributes, land use characteristics, multimodal public transit conditions, and bike-sharing self-loop intensity. We fill the existing gaps in the literature regarding the role of built environment characteristics and multimodal public transit in shaping the bike-sharing self-loop phenomenon. The primary contribution of this study lies in providing nuanced insights into the heterogeneous bike-sharing self-loop intensity at different

spatial scales, contributing to more effective policies to enhance integration between bike-sharing and public transit systems, and optimizing the bike-sharing operation efficiency.

The main findings can be summarized as follows: (1) The spatial spillover effect was significant at the street scale but not at the metro station scale, indicating distinct spatial interactions between these two scales in the bike-sharing self-loop phenomenon. (2) Land use emerged as the most critical determinant at both spatial scales, with residential areas showing the strongest positive impact on bike-sharing self-loop intensity, especially at the street scale. (3) Location attributes exhibited heterogeneous effects, with universities and transportation hubs promoting higher bike-sharing self-loop intensity at the metro station scale, while CBDs showed positive effects at the street scale, highlighting the distinct roles of these areas in shaping bike-sharing usage patterns. (4) Bus numbers and metro station ridership significantly influenced bike-sharing self-loop intensity, emphasizing the importance of bike-sharing as a feeder mode for multimodal public transit systems.

Based on these findings, we propose several targeted policy recommendations. At high-ridership metro stations near universities and residential areas, establishing dedicated cycling lanes is crucial for improving first- and last-mile connectivity. In transportation hub areas, scientifically designated geo-fencing for bike-sharing parking during peak hours and enhanced wayfinding can support better integration and reduce operational inefficiencies. For areas with fewer bus stations, prioritizing bike-sharing deployment as a complement to the bus network can serve to bridge gaps in transit access, thereby extending the reach of public transportation to underserved areas. By aligning bike-sharing resources with both land use and transportation conditions, these targeted interventions can enhance the overall efficiency of urban

mobility systems.

Despite the valuable insights provided by this research, several limitations should be acknowledged. First, the scope of this analysis is limited to a single city, which might restrict the applicability of our findings to other urban environments that differ in socioeconomic or infrastructural conditions. Second, this study primarily focuses on land use and transportation conditions, while other influential factors, such as real-time user behavior and economic incentives were not included. These factors could further refine our understanding of bike-sharing usage patterns. Future research should aim to expand the geographical scope of the analysis, include more influencing variables, and apply longitudinal methods to capture temporal dynamics in the bike-sharing self-loop phenomenon. Additionally, incorporating counterfactual inference could offer a more robust understanding of the underlying mechanisms driving bike-sharing self-loop intensity, ultimately leading to more precise and effective policy interventions.

Appendix

Table A1 presents the results of the Moran test. The Global Moran's I value of 0.691 at the metro station scale and 0.518 at the street scale demonstrates significant spatial autocorrelation in bike-sharing self-loop intensity, thereby underscoring the necessity of applying spatial econometric models for the analysis. Table A2 shows the collinearity test results. The Variance Inflation Factor (VIF) values for all variables are well below the threshold of 10, indicating no multicollinearity among the variables at either the metro station or street scales.

Table A1. Moran test results of bike-sharing self-loop intensity.

Scale	Metro station	Street
Global Moran's I	0.691	0.518

Table A2. Collinearity test of influencing factors on bike-sharing self-loop intensity.

Variable	Metro station scale		Street scale	
	VIF	1/VIF	VIF	1/VIF
Average age	2.410	0.414	2.410	0.415
Fixed occupations percentage	1.410	0.711	1.540	0.648
Non-resident population percentage	1.580	0.634	1.670	0.600
Average household car ownership	1.360	0.735	1.460	0.685
Work POI intensity	2.820	0.354	2.610	0.383
Residential POI intensity	6.050	0.165	4.620	0.217
Commercial POI intensity	4.130	0.242	4.700	0.213
Central area dummy	2.060	0.485	1.540	0.647
University area dummy	1.240	0.808	1.160	0.861
Central business district area dummy	1.260	0.791	1.300	0.768
Transportation hub area dummy	1.330	0.753	1.120	0.891
Metro station ridership	1.720	0.582		
Bus station number	1.500	0.669	1.940	0.514
Mean VIF	2.221		2.173	

References

BACH, P., KURZ, M., CHERNOZHUKOV, V., SPINDLER, M. & KLAASSEN, S. 2024. DoubleML : An Object-Oriented Implementation of Double Machine Learning in R. *Journal of Statistical Software,* 108.

BIELIŃSKI, T., KWAPISZ, A. & WAŻNA, A. 2021. Electric bike-sharing services mode substitution for driving, public transit, and cycling. *Transportation Research Part D: Transport and Environment,* 96, 102883.

BRUZZONE, F., SCORRANO, M. & NOCERA, S. 2021. The combination of e-bike-sharing and demand-responsive transport systems in rural areas: A case study of Velenje. *Research in Transportation Business & Management,* 40, 100570.

CAGGIANI, L., COLOVIC, A. & OTTOMANELLI, M. 2020. An equality-based model for bike-sharing stations location in bicycle-public transport multimodal mobility. *Transportation Research Part A: Policy and Practice,* 140, 251-265.

CHEN, Q., FU, C., ZHU, N., MA, S. & HE, Q.-C. 2023. A target-based optimization model for bike-sharing systems: From the perspective of service efficiency and equity. *Transportation Research Part B: Methodological,* 167, 235-260.

CHENG, L., HUANG, J., JIN, T., CHEN, W., LI, A. & WITLOX, F. 2023. Comparison of station-based and free-floating bike-share systems as feeder modes to the metro. *Journal of Transport Geography,* 107, 103545.

CHENG, Y., WANG, J. & WANG, Y. 2021. A user-based bike rebalancing strategy for free-floating bike sharing systems: A bidding model. *Transportation Research Part E: Logistics and Transportation Review,* 154, 102438.

CHU, J., DUAN, Y., YANG, X. & WANG, L. 2021. The last mile matters: Impact of dockless bike

sharing on subway housing price premium. *Management Science,* 67, 297-316.

DIAO, M., SONG, K., SHI, S., ZHU, Y. & LIU, B. 2023. The environmental benefits of dockless bike sharing systems for commuting trips. *Transportation Research Part D: Transport and Environment,* 124, 103959.

DURAN-RODAS, D., WRIGHT, B., PEREIRA, F. C. & WULFHORST, G. 2021. Demand And/oR Equity (DARE) method for planning bike-sharing. *Transportation Research Part D: Transport and Environment,* 97, 102914.

EREN, E. & UZ, V. E. 2020. A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society,* 54, 101882.

GAO, K., YANG, Y., LI, A. & QU, X. 2021. Spatial heterogeneity in distance decay of using bike sharing: An empirical large-scale analysis in Shanghai. *Transportation Research Part D: Transport and Environment,* 94, 102814.

GODAVARTHY, R., MATTSON, J. & HOUGH, J. 2022. Impact of bike share on transit ridership in a smaller city with a university-oriented bike share program. *Journal of Public Transportation,* 24, 100015.

GUO, Y. & HE, S. Y. 2020. Built environment effects on the integration of dockless bike-sharing and the metro. *Transportation Research Part D: Transport and Environment,* 83, 102335.

GUO, Y. & HE, S. Y. 2021. The role of objective and perceived built environments in affecting dockless bike-sharing as a feeder mode choice of metro commuting. *Transportation Research Part A: Policy and Practice,* 149, 377-396.

HOSFORD, K., YANAGAWA, C., LORE, M. & WINTERS, M. 2024. Effects of Mobi's equity initiatives on public bike share access and use. *Transportation Research Part D: Transport and Environment,* 131, 104223.

HUANG, G. & XU, D. 2023. The last mile matters: Impact of dockless bike-sharing services on traffic congestion. *Transportation Research Part D: Transport and Environment,* 121, 103836.

JABER, A. & CSONKA, B. 2023. How do land use, built environment and transportation facilities affect bike-sharing trip destinations? *Promet-Traffic & Transportation,* 35, 119-132.

JI, S., HEINEN, E. & WANG, Y. 2023. Non-linear effects of street patterns and land use on the bike-share usage. *Transportation Research Part D: Transport and Environment,* 116, 103630.

JIN, H., LIU, S., SO, K. C. & WANG, K. 2022a. Dynamic incentive schemes for managing dockless bike-sharing systems. *Transportation Research Part C: Emerging Technologies,* 136, 103527.

JIN, Y., RUIZ, C. & LIAO, H. 2022b. A simulation framework for optimizing bike rebalancing and maintenance in large-scale bike-sharing systems. *Simulation Modelling Practice and Theory,* 115, 102422.

KIM, M. & CHO, G.-H. 2021. Analysis on bike-share ridership for origin-destination pairs: Effects of public transit route characteristics and land-use patterns. *Journal of Transport Geography,* 93, 103047.

KONG, H., JIN, S. T. & SUI, D. Z. 2020. Deciphering the relationship between bike sharing and public transit: Modal substitution, integration, and complementation. *Transportation Research Part D: Transport and Environment,* 85, 102392.

LI, C., CHEN, G. & WANG, S. 2024a. Urban mobility resilience under heat extremes: Evidence from bike-sharing travel in New York. *Travel Behaviour and Society,* 37, 100821.

LI, W., CHEN, S., DONG, J. & WU, J. 2021. Exploring the spatial variations of transfer distances between dockless bike-sharing systems and metros. *Journal of Transport Geography,* 92, 103032.

LI, X., XU, Y., ZHANG, X., SHI, W., YUE, Y. & LI, Q. 2023. Improving short-term bike sharing demand forecast through an irregular convolutional neural network. *Transportation Research Part C: Emerging Technologies,* 147, 103984.

LI, Y., WANG, S., TANG, J. H. C. G., PENG, Z. & ZHUGE, C. 2024b. Public attention and attitudes towards bike-sharing in China: A text mining approach. *Transportation Research Part D: Transport and Environment,* 134, 104348.

LIU, L., KONG, H., LIU, T. & MA, X. 2022. Mode Choice between Bus and Bike-Sharing for the Last-Mile Connection to Urban Rail Transit. *Journal of Transportation Engineering, Part A: Systems,* 148, 04022017.

LIU, L., LEE, J. & MILLER, H. J. 2024. Evaluating accessibility benefits and ridership of bike-transit integration through a social equity lens. *Computers, Environment and Urban Systems,* 112, 102150.

LIU, X. C., TAYLOR, J., PORTER, R. J. & WEI, R. 2018. Using trajectory data to explore roadway characterization for bike-share network. *Journal of Intelligent Transportation Systems,* 22, 530-546.

LUO, H., ZHAO, F., CHEN, W.-Q. & CAI, H. 2020. Optimizing bike sharing systems from the life cycle greenhouse gas emissions perspective. *Transportation Research Part C: Emerging Technologies,* 117, 102705.

MIX, R., HURTUBIA, R. & RAVEAU, S. 2022. Optimal location of bike-sharing stations: A built environment and accessibility approach. *Transportation Research Part A: Policy and Practice,* 160, 126-142.

PEKDEMIR, M. I., ALTINTASI, O. & OZEN, M. 2024. Assessing the Impact of Public Transportation, Bicycle Infrastructure, and Land Use Parameters on a Small-Scale Bike-Sharing System: A Case Study of Izmir, Türkiye. *Sustainable Cities and Society,* 101, 105085.

SALTYKOVA, K., MA, X., YAO, L. & KONG, H. 2022. Environmental impact assessment of bike-sharing considering the modal shift from public transit. *Transportation Research Part D: Transport and Environment,* 105, 103238.

SEMENOVA, V. & CHERNOZHUKOV, V. 2020. Debiased machine learning of conditional average treatment effects and other causal functions. *The Econometrics Journal,* 24**,** 264-289.

SHI, Z., WANG, J., LIU, K., LIU, Y. & HE, M. 2024. Exploring the usage efficiency of electric bike-sharing from a spatial–temporal perspective. *Transportation Research Part D: Transport and Environment,* 129**,** 104139.

SONG, Y., ZHANG, L., LUO, K., WANG, C., YU, C., SHEN, Y. & YU, Q. 2024. Self-loop analysis based on dockless bike-sharing system via bike mobility chain: empirical evidence from Shanghai. *Transportation*.

SUN, S. & ERTZ, M. 2022. Can shared micromobility programs reduce greenhouse gas emissions: Evidence from urban transportation big data. *Sustainable Cities and Society,* 85**,** 104045.

WANG, B., GUO, Y., CHEN, F. & TANG, F. 2024. The impact of the social-built environment on the inequity of bike-sharing use: A case study of Divvy system in Chicago. *Travel Behaviour and Society,* 37**,** 100873.

WANG, X., CHENG, Z., TRéPANIER, M. & SUN, L. 2021. Modeling bike-sharing demand using a regression model with spatially varying coefficients. *Journal of Transport Geography,* 93**,** 103059.

WU, C. & KIM, I. 2020. Analyzing the structural properties of bike-sharing networks: Evidence from the United States, Canada, and China. *Transportation Research Part A: Policy and Practice,* 140**,** 52-71.

XIN, R., YANG, J., AI, B., DING, L., LI, T. & ZHU, R. 2023. Spatiotemporal analysis of bike mobility chain: A new perspective on mobility pattern discovery in urban bike-sharing system. *Journal of Transport Geography,* 109**,** 103606.

XING, Y., WANG, K. & LU, J. J. 2020. Exploring travel patterns and trip purposes of dockless bike-sharing by analyzing massive bike-sharing data in Shanghai, China. *Journal of Transport Geography,* 87**,** 102787.

YAN, Y. & CHEN, Q. 2024. Spatial heterogeneity and nonlinearity study of bike-sharing to subway connections from the perspective of built environment. *Sustainable Cities and Society,* 114**,** 105766.

YU, C., DONG, W., LIU, Y., YANG, C. & YUAN, Q. 2024. Rethinking bus ridership dynamics: Examining nonlinear effects of determinants on bus ridership changes using city-level panel data from 2010 to 2019. *Transport Policy,* 151**,** 85-100.

YU, Q., LI, W., YANG, D. & XIE, Y. 2020. Policy zoning for efficient land utilization based on spatio-temporal integration between the bicycle-sharing service and the metro transit. *Sustainability,* 13**,** 141.

ZHANG, Y., SHAO, Y., BI, H., AOYONG, L. & YE, Z. 2023. Bike-sharing systems rebalancing considering redistribution proportions: A user-based repositioning approach. *Physica A: Statistical Mechanics and its Applications,* 610**,** 128409.

ZHAO, J., FAN, W. & ZHAI, X. 2020. Identification of land-use characteristics using bicycle sharing data: A deep learning approach. *Journal of Transport Geography,* 82**,** 102562.

ZHAO, S., ZHAO, K., XIA, Y. & JIA, W. 2022. Hyper-clustering enhanced spatio-temporal deep learning for traffic and demand prediction in bike-sharing systems. *Information Sciences,* 612**,** 626-637.

ZHENG, H., ZHANG, K., NIE, Y., YAN, P. & QU, Y. 2024. How Many Are Too Many? Analyzing Dockless Bike-Sharing Systems with a Parsimonious Model. *Transportation Science,* 58**,** 152-175.

ZHOU, X. & XIE, Y. 2020. Heterogeneous Treatment Effects in the Presence of Self-selection: A Propensity Score Perspective. *Sociological Methodology,* 50**,** 350-385.

ZHOU, X., ZHAO, Z., FU, W., HUANG, Z., YAO, Y., HUANG, Y. & ZHANG, Y. 2024. The impact of heterogeneous accessibility to metro stations on land use changes in a bike-sharing context. *Journal of Transport Geography,* 121**,** 104019.

ZHUANG, D., JIN, J. G., SHEN, Y. & JIANG, W. 2021. Understanding the bike sharing travel demand and cycle lane network: The case of Shanghai. *International Journal of Sustainable Transportation,* 15**,** 111-123.